

Gerald Neumann, Ute Recker-Hamm:

Abstract zum Beitrag "Zur elektronischen Vernetzung von Wörterbüchern: Aspekte, Voraussetzungen, Qualitätsmerkmale"

Die rasante Entwicklung und allgemeine Verbreitung der elektronischen Medien hat zur Entstehung von Vorstellungen geführt, die auch die Erwartungen an elektronische wissenschaftliche Wörterbücher prägen. Das Internet mit seinen Möglichkeiten der variablen Darstellung und Vernetzung von Informationen sowie mit seinen neuartigen Suchfunktionen eröffnet eine Fülle von Zugriffs- und Nutzungsmöglichkeiten, die bislang von den vorhandenen Wörterbuchressourcen nur ansatzweise realisiert worden sind. Dass nicht alles, was technisch möglich ist, für ein bestimmtes Wörterbuch mit seinem spezifischen Darstellungszweck und -gegenstand oder auch für Sprachwörterbücher im Allgemeinen sinnvoll ist, soll hier nicht in Frage gestellt werden. In unserem Beitrag, der sich als Arbeits- und Diskussionsvorschlag versteht, wird zunächst eine Bestandsaufnahme des bislang realisierten Funktions- und Vernetzungsumfangs anhand von fünf ausgewählten, im Internet verfügbaren digitalen Wörterbuch-Ressourcen vorgenommen. Untersucht wird erstens das Trierer Wörterbuchnetz (www.woerterbuchnetz.de), das einen Verbund aus verschiedenen, retrodigitalisierten Wörterbüchern bildet, darunter das 'Deutsche Wörterbuch (DWB)' der Brüder Grimm, die Mittelhochdeutschen Wörterbücher von Benecke, Müller, Zarncke und Lexer sowie eine Reihe von Dialektwörterbüchern und das 'Goethe-Wörterbuch (GWB)'; zweitens das in Bearbeitung befindliche neue 'Mittelhochdeutsche Wörterbuch (MWB)' (www.mhdwb-online.de); drittens das 'Deutsche Rechtswörterbuch (DRW)' (<http://drw-www.adw.uni-heidelberg.de/drw/>), das sowohl retrodigitalisierte als auch neu erarbeitete Artikelstrecken enthält; viertens das 'Digitale Wörterbuch der deutschen Sprache des 20. Jahrhunderts' (www.dwds.de), das das retrodigitalisierte 'Wörterbuch der deutschen Gegenwartssprache (WDG)' mit selbstständigen Corpora und mit durch computerlinguistische Methoden erhobenen Zusatzinformationen kombiniert; und fünftens schließlich das 'Etymologische Wörterbuch des Deutschen (EtymWB)' von Wolfgang Pfeifer, das ebenfalls retrodigitalisiert wurde und uns in einer ersten Rohfassung vorlag.

Die Bestandsaufnahme wird mit Hilfe einer detaillierten Checkliste durchgeführt, die die folgenden zusammengefassten Gesichtspunkte berücksichtigt:

1. Umfang und Inhalt der Ressource, Zugangsart, Navigation, Ansichten, Zitierfähigkeit, Verhältnis zur Druckfassung, Verfügbarkeit und Performanz
2. Realisierung von expliziten oder impliziten Verweisen innerhalb der Ressource. Zu den impliziten Verweisen werden z.B. registerartige Strukturen gerechnet und die Visualisierung von Gegenverweisen.
3. Verlinkung der Wörterbuchkomponente mit weiteren internen Ressourcen, wie Quellenverzeichnis, Quellencorpus, Belegarchiv usw.
4. Verlinkung mit externen Online-Angeboten, z.B. anderen Wörterbüchern
5. Suchmöglichkeiten, Darstellung der Suchergebnisse und Fehlerbehandlung

6. Adressierbarkeit der Bestandteile, d.h. wie können z.B. einzelne Artikel, Einträge des Quellenverzeichnisses usw. von außen angesprochen werden.
7. Standardisierung und Schnittstellen für nicht Browser-bezogenen Zugriff wie z.B. Webservices

Neben beeindruckenden individuellen Stärken der einzelnen Wörterbuch-Angebote zeigt die Bestandsaufnahme eine überraschende gemeinsame Schwäche fast aller untersuchten Ressourcen im Bereich der Suchmöglichkeiten: Während in allen betrachteten Wörterbüchern die Suche nach Stichwörtern selbstverständlich ist und die Hauptzugriffsstruktur darstellt, steht (mit Ausnahme des EtymWB) in keinem eine Volltextsuche oder eine Suche in Feldern (z.B. nur im lexikographischen Kommentar) zur Verfügung. Dieser Befund ist umso erstaunlicher, als die Suchmöglichkeiten und die sich daraus neu ergebenden Zugriffsmöglichkeiten im Allgemeinen als besondere Stärke elektronischer Medien gelten.

Was die Vernetzung der elektronischen Wörterbücher untereinander angeht, die insbesondere wegen ihrer vielschichtigen inhaltlichen Überlappungen ein besonderes Potential aufweist¹, so ist zu beobachten, dass es bislang nur recht spärlich ausgenutzt wird: Das DRW und das MWB verweisen systematisch auf korrespondierende Artikel in den Mittelhochdeutschen Wörterbüchern des Trierer Wörterbuchnetzes, das DRW zusätzlich auf das DWB. Das EtymWB ist mit dem WDG verlinkt. Für die fehlenden Verknüpfungen scheinen zwei Ursachen im Vordergrund zu stehen: Erstens bietet keine Ressource eine komfortable oder standardisierte Möglichkeit, die nötigen Linkadressen systematisch zu ermitteln, so dass die Verlinkung ohne Mithilfe des Zielwörterbuchs erfolgen könnte. Zweitens erfordert die qualitativ hochwertige Verlinkung ein nicht unbeträchtliches Maß an manuellem Eingreifen, wie die verschiedenen manuellen bzw. algorithmischen Verlinkungsstrategien zeigen, die innerhalb des Trierer Wörterbuchnetzes zur Anwendung kamen.

Die Checkliste sowie die zusammenfassenden Beobachtungen zur Suche und zur Vernetzung sollen Anregungen für die Diskussion über digitale Wörterbücher und ihre zukünftige Weiterentwicklung bieten.

¹ Vgl. dazu *Das Mittelhochdeutsche Wörterbuch: Beleglexikographische Konzeption, EDV, Vernetzungspotentiale*. In: *Lexicographica* 23 (2007), S. 77-95.